City Research Online

City, University of London Institutional Repository

Citation: Yarrow, K. (2017). Assessing Risk Factors for Athletic Excellence. Progress in Brain Research, doi: 10.1016/bs.pbr.2016.11.008

This is the accepted version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/16376/

Link to published version: https://doi.org/10.1016/bs.pbr.2016.11.008

Copyright: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

Reuse: Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

City Research Online: http://openaccess.city.ac.uk/ publications@city.ac.uk

Assessing Risk Factors for Athletic Excellence

Kielan Yarrow

Cognitive Neuroscience Research Unit, Department of Psychology,

City, University of London,

London, U.K.

Abstract

Hardy et al. present a qualitative interview study to identify differences between elite and super-elite athletes, but their approach closely parallels the case-control methodology developed and discussed primarily within the epidemiological literature. The strengths and weaknesses of this approach may be less familiar to a psychological or sports-science audience. Here, I comment on this methodological choice, and the implications it has for the robustness and value of the inferences that are drawn. I conclude that while Hardy et al. have made valuable progress in dissecting the factors that generate elite athletic performance, their work is best considered as exploratory, and would benefit greatly from confirmation via a prospective cohort approach.

Keywords

Case-control, cohort, elite, athlete, retrospective.

In their thought-provoking paper, Hardy et al. use a qualitative analysis of interviews with (matched) elite and super-elite athletes in order to identify putative life events, and also personality and motivational factors, that differed between these groups. They particularly highlight eight features that are seen in a much higher proportion of the super-elite athletes compared to their elite colleagues. In most cases, I found the observed differences both intriguing and well evidenced by the supporting quotations.[1] In what follows, I examine their basic research methodology at a fairly abstract level. This approach leads me to conclude that the data presented in this target article are intriguing and suggestive, but some way away from being conclusive.

Hardy et al. use an interview methodology that generates rich qualitative data, but their approach is nonetheless recognisable as a *case-control* study (with the super-elite multi-medal-winning athletes as the cases, and the elite athletes that never medalled at the World Championships or Olympic games as the controls). The defining feature of a case-control study is that cases are identified, then matched with controls, and the two groups are compared retrospectively in order to determine the relative frequencies of exposure to one or more hypothesised risk factors. In Hardy et al.'s paper, risk factors are instead called variables, reflecting the psychological rather than medical tradition in which they work. Case-control approaches rarely receive explicit mention within Psychology, but the strengths and weaknesses of this kind of study are well known from the epidemiological literature (e.g. Coggon, Barker & Rose, 2003; Sedgewick, 2015).

In general, case-control studies are ranked below randomised controlled trials and cohort studies in terms of their methodological persuasiveness, but of course they are much easier to run. That said, for the kinds of variables considered here, it's unlikely we could ever envisage a true experiment, with random allocation to conditions and so forth. However, a cohort study might be feasible. I certainly don't wish to belittle what has been accomplished, but we should be realistic about its limitations. Here, I will consider three aspects of Hardy et al.'s work in a bit more detail: The robustness of statistical inference, the reliability of retrospective data, and the appropriateness of the controls. I won't comment in detail on the essentially correlational nature of the data, as I think Hardy et al. are already fairly clear in their caveats on that front, but this should of course also be borne in mind.

1. Robustness of statistical inference

[1] The exception, for me, was the ability to perform under pressure, which in itself seemed rather tautological, and when interpreted via the psychodynamic concept of counterphobia appeared a slightly bewildering leap of faith in light of what the athletes had actually said. Still, seven out of eight is pretty good.

Hardy et al. tread a slightly opaque path between a qualitative and quantitative methodology, in which they identify risk factors, and the frequencies with which they occur in both super-elite and elite athletes, but don't present the inferential stats that might typically accompany these descriptions. Thankfully, even the least discriminating of their differences (the "need for success", found in 13 of 16 super elites but only 6 of 16 elites) is in fact conventionally significant ($\chi^2_{[1]} = 6.35$, $p = 0.0118$). However, this issue does deserve a little further thought. In a typical case-control study, the risk factors of interest are identified in advance, legitimising an *a priori* statistical comparison. Here, the variables are generated from the qualitative analysis of an interview. Although themes are clearly identified in advance, the actual risk factors within each theme emerge only after engagement with the data. Hence we are in the realm of *post hoc* analysis, but with an ill-defined number of variables to consider for the correction of familywise error rates. I certainly don't think that this invalidates Hardy et al.'s findings, but it should lead us to weigh the study as an exploratory, rather than confirmatory, exercise. I look forward to subsequent confirmatory work.

2. The reliability of retrospective data.

Hardy et al. allude to the fragility of human memory in their methods section, and make efforts to confirm their data via interviews with coaches and parents. This is laudable. What worries me slightly is the potential for a kind of groupthink in the narratives that athletes, coaches and their families develop to account for life outcomes. I don't doubt that the super-elite athletes did experience negative early life events around the time of a significant (re)commitment to their sport, but I wonder how pondering one's past from the position of having to explain subsequent tremendous success might affect the interpretation and rehearsal of teenage memories. And could there be some post-hoc rationalisation by the elites, when contemplating the role that their personality, motivation and lifestyle played in their failure to reach the highest heights? Did they fail because they "wanted it" but didn't "need it", or did they conclude that they didn't "need it" because they failed? These are the sorts of concerns that elevate prospective longitudinal studies over retrospective ones.

3. The appropriateness of the controls.

This is an important issue in case-control methodology, and one which, in fairness, Hardy et al. score well on. The super-elites were matched with athletes of the same sex/sport/age/era who had been recognised as worthy investments by a national sporting body but never medalled at major events. That seems both clear and sensible to me. One possible issue is that we are not told how these controls were actually selected. Ideally, they would have been picked completely at random from all those who met the matching criteria, and indeed this may have been the case. The worry would be

that (unconsciously or otherwise) the researchers were directed to controls that exhibited the sorts of traits and histories that coaches *believe* undermined their success. Note that I am not accusing the study team of any such bias, intentional or unintentional. I am simply highlighting some of the subtle confounds that can occur in even well-conducted research.

To summarise then: Hardy et al. use what is essentially a case-control methodology to identify factors that differentiate super-elite athletes from elite athletes, and come up with some interesting potential risk factors for Olympic gold, such as escaping an early life trauma through sport, and the ruthless pursuit of athletic goals. The case-control methodology is well established, and has helped us to move forward with many important questions. For example, a case-control study provided the first evidence that smoking causes cancer (Doll & Hill, 1950). However, it is important not to get carried away with the allure of this design, which might be falsely enhanced when supplemented by powerfully persuasive anecdotes (as here). The tobacco companies were defeated by cohort, not case-control, studies. Ultimately, we will need additional evidence to corroborate what has been hinted at by Hardy et al. about the factors that generate athletic greatness. Fortunately, with their meticulous approach, they seem just the people to provide it.

References

Coggon, D., Barker, D., & Rose, G. (2003). *Epidemiology for the Uninitiated (5th Edition)*. John Wiley & Sons.

Doll R, Hill AB (1950). Smoking and carcinoma of the lung; preliminary report. *BMJ, 2* (4682), 739–48.

Sedgwick, P. (2015). Bias in observational study designs: case-control studies. *BMJ*, *350*, h560.